Plan Overview

A Data Management Plan created using DMPonline

Title: SSpaceGX: Social Space and Nature Conservation in the Gerêz-Xurés Transboundary Biosphere Reserve (Portugal/Spain)

Creator: Raimundo Elias Gomez

Affiliation: University of Porto (Universidade do Porto)

Template: DCC Template

Project abstract:

The project's main objective is to create a multidimensional model of the socio-environmental dynamics of the Gerêz-Xurés Transboundary Biosphere Reserve (located on the borders of Portugal and Spain). For this purpose, we will analyze the social space of nature conservation in this Biosphere Reserve over the last decades to the present (years 2000 to 2023). The specific objectives take on multiple empirical aspects of the social space, which will be analyzed quantitatively and qualitatively. These aspects encompass: (I) the characterization of the main involved agents, including their capitals, strategies, and positions; (II) the social-environmental transformations caused by districts, national and transnational, development, and nature conservation programs; and (III) the evolution of household living conditions within internal jurisdictions (parishes from six municipalities in Portugal and six municipalities in Spain).

ID: 103692

Start date: 10-08-2023

End date: 10-08-2025

Last modified: 01-09-2022

Copyright information:

SSpaceGX: Social Space and Nature Conservation in the Gerêz-Xurés Transboundary Biosphere Reserve (Portugal/Spain)

Data Collection

What data will you collect or create?

We will work with multiple types of data (distinct in terms of collection form, storage format, and structuring mode).

- **Background data**: this type of data includes previously done research and reports on people, the environment, and local practices. In short, a collection of virtual files that will not exceed 100 GB (in .pdf, .doc, .rtf, .flv, .avi, or .wmv format). Those files with an open access license can be saved and shared in the author's account on [GitHub](#) and [Zenodo](#). For background data, bibliographic searches will be conducted in libraries, on the Internet, and in newspaper archives associated with the biosphere reserve. These searches will be conducted in conjunction with the ethnographic work, which will include interviews, observations, conversations, and workshops with members of private and civil society organizations.
- **Spatial data**: the geo-referenced data about the biosphere reserve for the years 2001, 2011, and 2021, involves mostly satellite imagery and map bases created or downloaded from public repositories (in .tiff, .jpg, .json, .sh, or .ogc format). Those files with an open access license can be saved and shared in [the author's account on GitHub](#) and [Zenodo](#). The geo-referenced data about the biosphere reserve for the years 2001, 2011, and 2021 will be downloaded from international and public agencies such as the [Copernicus Open Access Hub](#), the [Unites States Geological Survey](#), the [National Geographic Institute of Spain](#), and the [National Geographic Information System of Portugal](#)
- **Statistical data**: This type of data involves primarily statistical microdata from national censuses and households' permanent surveys of Portugal and Spain about the Gerêz-Xurés Transboundary Biosphere Reserve. Basically, a collection of six statistical matrices, one for each municipality within the Biosphere Reserve for the years 2001, 2011, and 2021 (in .csv and, .sav format). Files with an open access license and a complete anonimization of confidential data can be saved and shared in [the author's account on GitHub](#) and [Zenodo](#). The statistical microdata from national censuses and households' permanent surveys of Portugal and Spain will be mainly collected from the [National Statistics Institute of Portugal](#), from the [National Statistics Institute of Spain](#), and from the [Integrated Public Use Microdata Series](#).
- **Etnographic data**: That is, researches based on a qualitative approach to local inhabitants, especially through fieldwork and residence in the transition and buffer zones of the Gerêz-Xurés Transboundary Biosphere Reserve. In summary, include content in various audio, video, and text templates (in .jpg, .wmv, .flv, .rtf, or .doc format). Only files that do not violate the privacy of local residents and have their permissions will be saved and shared in [the author's account on GitHub](#) and [Zenodo](#). this type of data requires long periods of residence in the Transition and Buffer Zones of the Gerêz-Xurés Transboundary Biosphere Reserve (areas where human populations and productive activities are permitted in accordance with UNESCO Biosphere Reserve guidelines). Face-to-face interaction with neighbors and members of private and civil society organizations is the primary source of ethnographic data; above all, by participating in public and daily activities with locals.

The statistical matrices, the spatial and satellite imagery set of layers, the etnographic data, as well as the socio-environmental model in course of completion, will be stored at the author GitHub repository in https://github.com/debianalt and [Zenodo](#) (a multidisciplinary open access repository developed under the European OpenAIRE program and operated by CERN [European Organization for Nuclear Research]). GitHub is an online coding development platform used for storing, tracking, and collaborating on open-source projects. It allows researchers to upload their own code files and join forces with other researchers working on the same or similar projects. The statistical and spatial analysis will be done mainly with "R", a free software environment for statistical computing and graphics (https://www.r-project.org). "R" will be utilized in conjunction with GitHub (and GitHub have a complete integration with [Zenodo](#)). So, the final advances can be checked from GitHub, in fact, the complete research can be consulted, replicated, and tested by downloading the analysis chains.

How will the data be collected or created?

1) *Background data*: Bibliographic searches will be conducted in libraries, on the Internet, and in newspaper archives associated with the biosphere reserve. This search will be conducted in conjunction with the ethnographic work, which will include interviews, observations, conversations, and workshops with members of private and civil society organizations.

2) *Spatial data*: The geo-referenced data about the biosphere reserve for the years 2001, 2011, and 2021 will be downloaded from international and public agencies such as the *Copernicus Open Access Hub* (https://scihub.copernicus.eu/dhus/#/home), the *Unites States Geological Survey* (https://earthexplorer.usgs.gov/), the *National Geographic Institute of Spain* (https://www.ign.es/web/ign/portal), and the *National Geographic Information System of Portugal* (https://snig.dgterritorio.gov.pt/)

3) *Statistical data*: The statistical microdata from national censuses and households' permanent surveys of Portugal and Spain about the Gerêz-Xurés Transboundary Biosphere Reserve will be collected from the *National Statistics Institute of Portugal* (https://www.ine.pt/), from the *National Statistics Institute,* and from the *Integrated Public Use Microdata Series* (https://international.ipums.org/international/).

4) *Etnographic data*: this type of data requires long periods of residence in the Transition and Buffer Zones of the Gerêz-Xurés Transboundary Biosphere Reserve (areas where human populations and productive activities are permitted in accordance with

biosphere reserve guidelines). Face-to-face interaction with neighbors and members of private and civil society organizations is the primary source of ethnographic data; above all, by participating in public and daily activities with locals.

*The statistical matrices, the spatial and satellite imagery set of layers, the etnographic data, as well as the socio-environmental model in course of completion, will be stored at the author GitHub repository in https://github.com/debianalt and Zenodo.The usual structure used for the production and storage of models (and programs) on GitHub will be resumed. This structure includes the following folders: "data", "figures", "outputs", "papers", "reports", and "scripts". The "data" folder is subdivided into two others named "raw data" and "processed data". In each of the folders called "raw data" and "processed data" four more will be added, labeled as: "background data", "spatial data", "statitical data", and "etnographic data". The names of each stored file will include year, location, and responsible author/organization; and metadata adjusted to the recommendations of the "CESSDA Metadata Validator" (Consortium of European Social Science Data Archives) will be added to these files; so **the raw and processed data will be Findable, Accessible, Interoperable, and Reusable**. Finally, any information that could endanger confidentiality or other civil rights of the local population will not be shared or stored. Furthermore, the legal requirements of each of the data sources used will be respected.

Documentation and Metadata

What documentation and metadata will accompany the data?

Metadata will be created using *Dublin Core* Metadata Element. This platform and community of metadata professionals will be very helpful for the creation of specific descriptors for the research process.

- Data Documentatio Initiative (DDI) will be used for the description of statistical and geospatial techniques, fieldwork instruments, and social theory topics.
- More details will be added in an updated version of this document.

Ethics and Legal Compliance

How will you manage any ethical issues?

Data Privacy Impact Assessment (DPIA) is not required for this project. As a result, any emerging issue involving the Biosphere Reserve residents' privacy, confidentiality, or any other civil right will be discussed with the research team and, eventually, with the Ethics Committee of the Faculty of Arts at the University of Porto.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

The raw data is partly open access. Satellite imagery and background research can be freely shared and stored because they are typically open acces (if any of them have proprietary restrictions, they will be excluded from storage and sharing).

Statistical microdata matrices are subject to confidentiality and privacy restrictions. To be stored and shared, the variables that endanger these rights must first be removed from the matrices. The same is true for certain ethnographic data (photographs, videos, and recipes, for example), which require the explicit permission of their authors or creators to be stored and shared. If such data is discovered, written permission will be required to share it.

The processed data will be restricted during project development and for six months after project completion in order to pursue scientific publications. Afterwards, the processed data and analyzes will be shared freely according to a Creative Commons license.

Storage and Backup

How will the data be stored and backed up during the research?

The RAW data will be saved on the collector's working laptop and desktop computer.
Weekly backups of the processed data will be created. The processed data will be kept on the author's working laptop and external

disks for 5 years, anonymized and coded.

The processed data, once completely anonymised, will be shared with the research team through shared cloud storage accounts. Once the research is finished and after 6 months, the processed data will be shared publicly through the Principal Investigator GitHub Account.

How will you manage access and security?

The raw data will be accessible only to the Principal Investigator and his Supervisor. The other members of the project team will only have access to processed data that does not contain any confidential, or deanonymized information.

All laptops and computers used to analyze and/or temporarily deposit project data will be subject to security procedures. To ensure up-to-date security, all computers will be password protected, and all software will be fully licensed and frequently updated.

In the future, a specific research data repository will be chosen, and its description will be included in an updated version of this DMP.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

The majority of data processed (and anonymized) will be saved and preserved after the project's conclusion (data will be stored and backed up for 5 years after research completion). The spatial, statistical, or ethnographic data that are involved with the model of socio-environmental interdependencies around the GXTBR will be especially protected.

As previously stated, the databases will be housed in Zenodo, the geographical and statistical analysis chains in GitHub, and the resulting publications and presentations will be hosted in the University of Porto's open-access repository.

What is the long-term preservation plan for the dataset?

The processed data will be stored and backed up for 5 years after research completion. Nevertheless, the articles, presentations, and other files resulting from the research that are held in the University of Porto's open repository will be long term accessible.

Data Sharing

How will you share the data?

The metadata will be open on the Research Data Repository of the University of Porto. Data resulting from the research will be available upon request after project publications.

Are any restrictions on data sharing required?

The processed data (and anonymized) have no restrictions on data sharing.

Responsibilities and Resources

Who will be responsible for data management?

Gomez Raimundo Elias, the SSPaceGX PI, is the main responsible for the management of all data.

What resources will you require to deliver your plan?

It is necessary to have options for partial data deposit within the Research Data Repository of the University of Porto.
Furthermore, the following assets will be used during the project:
Hardware/devices: work desktop and laptop computers, personal laptop computers; USB flash drives; external disks; institutional servers; cameras; microphones; loudspeakers; smartphones; tablets.
Software: "R", Rstudio, SPAD, QGIS, and SPSS.
Networks: entities' local networks and Wi-Fi networks with access to the Internet.
Message exchanges: entities' email services, Google "GMail" emails, Skype.
Paper transmission channels: notes related to teams, meetings, participants names, etc.